

PATENT ABSTRACTS OF JAPAN

(11) Publication number : 06-012097

(43) Date of publication of application : 21.01.1994

(51) Int.Cl.

G10L 9/14
G10L 9/18

(21) Application number : 04-170895

(71) Applicant : NIPPON TELEGR & TELEPH CORP <NTT>

(22) Date of filing : 29.06.1992

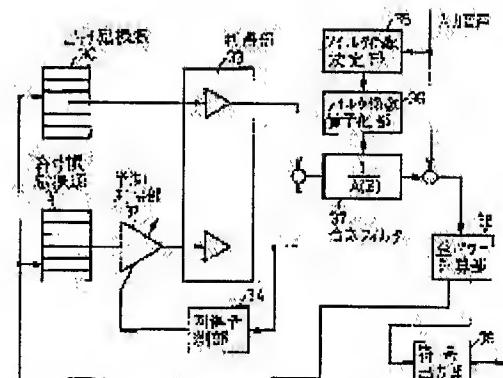
(72) Inventor : KATAOKA AKITOSHI
MORIYA TAKEHIRO

(54) METHOD AND DEVICE FOR PREDICTIVELY ENCODING VOICE

(57) Abstract:

PURPOSE: To provide the method and device for predictively encoding voice without transmitting any gain.

CONSTITUTION: In this predictive voice encoding method for encoding voices while utilizing the reproduction of voices by driving a synthesizing filter, to which a quantizing filter coefficient calculated based on the sampling of plural input voices is set, for the unit of a frame from two exciting sources corresponding to a time sequence vector component, the respective gains of the vector components at the two exciting sources are simultaneously obtained, and the gain to be imparted to the noise waveform vector components of a code book exciting source 31 of the next frame is predicted based on a residual waveform multiplied a gain selected by the selected noise waveform vector components of the code book exciting source 31 having the plural noise waveform vector components as one of the two exciting sources and a residual waveform stored before that so as to encode voices without transmitting the information of power.



LEGAL STATUS

[Date of request for examination]

24.02.1995

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 3087796

[Date of registration] 14.07.2000

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平6-12097

(43)公開日 平成6年(1994)1月21日

(51)Int.Cl.⁵

G 10 L 9/14

識別記号 庁内整理番号

G 8946-5H

F I

技術表示箇所

J 8946-5H

9/18

E 8946-5H

審査請求 未請求 請求項の数4(全6頁)

(21)出願番号

特願平4-170895

(22)出願日

平成4年(1992)6月29日

(71)出願人 000004226

日本電信電話株式会社

東京都千代田区内幸町一丁目1番6号

(72)発明者 片岡 章俊

東京都千代田区内幸町1丁目1番6号 日

本電信電話株式会社内

(72)発明者 守谷 健弘

東京都千代田区内幸町1丁目1番6号 日

本電信電話株式会社内

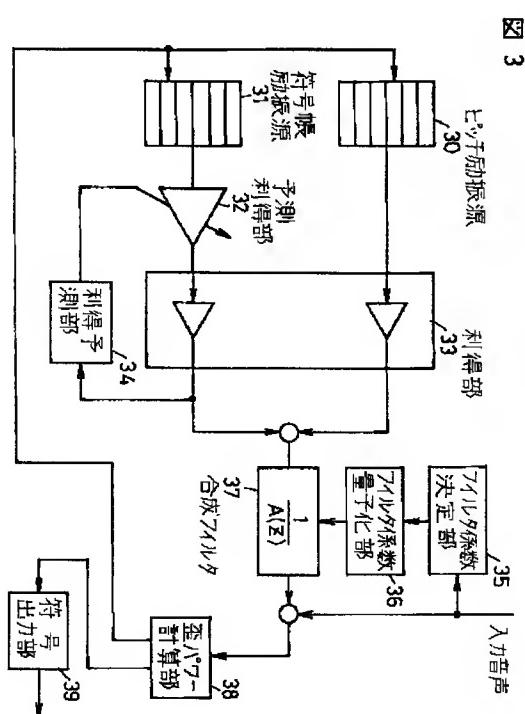
(74)代理人 弁理士 草野 卓

(54)【発明の名称】 音声の予測符号化方法および装置

(57)【要約】

【目的】利得を伝送することはしない音声予測符号化方法および装置を提供する。

【構成】入力音声を複数サンプリングしたものに基づいて計算された量子化フィルタ係数の設定された合成フィルタを2個の励振源から時系列ベクトル成分によりフレーム単位毎に駆動して音声を再生することを利用して音声を符号化する音声の予測符号化方法において、2個の励振源のベクトル成分に対してそれぞれの利得を同時に求め、2個の励振源の内の一方である複数の雜音波形ベクトル成分を有する符号帳励振源31の選択された雜音波形ベクトル成分に選択された利得を乗じた残差波形とそれ以前に蓄積されている残差波形とに基づいて次のフレームの符号帳励振源31の雜音波形ベクトル成分を与える利得を予測することにより、パワーの情報を伝送することなく音声を符号化する音声の予測符号化方法。



【特許請求の範囲】

【請求項1】 入力音声を複数サンプリングしたものに基づいて計算された量子化フィルタ係数の設定された合成フィルタを2個の励振源から時系列ベクトル成分によりフレーム単位毎に駆動して音声を再生することを利用して音声を符号化する音声の予測符号化方法において、2個の励振源のベクトル成分に対してそれぞれの利得を同時に求め、2個の励振源の内的一方である複数の雑音波形ベクトル成分を有する符号帳励振源の選択された雑音波形ベクトル成分に選択された利得を乗じた残差波形とそれ以前に蓄積されている残差波形に基づいて次のフレームの符号帳励振源の雑音波形ベクトル成分に与える利得を予測することにより、パワーの情報を伝送することなく音声を符号化する音声の予測符号化方法。

【請求項2】 入力音声を複数サンプリングしたものに基づいて計算された量子化フィルタ係数の設定された合成フィルタを具備し、複数のピッチ周期成分を有するピッチ励振源と複数の雑音波形ベクトル成分を有する符号帳励振源より成る2個の励振源からフレーム単位毎に時系列ベクトル成分により合成フィルタを駆動して音声を再生することを利用して音声を符号化する音声の予測符号化装置において、2個の励振源およびのベクトル成分それぞれに対して利得を同時に乗算する利得部を具備し、合成音声と入力音声との間の差である歪を最小にする様に両励振源中の各励振ベクトル成分を選択すると共に利得部の利得を設定する歪パワー計算部を具備し、符号帳励振源の選択された雑音波形ベクトル成分に選択された利得を乗じた残差波形とそれ以前に蓄積されている残差波形に基づいて次のフレームの符号帳の雑音波形ベクトル成分に与える利得を予測する利得予測部を具備することを特徴とする音声の予測符号化装置。

【請求項3】 請求項2に記載される音声の予測符号化装置において、利得部はピッチ励振源用利得部および符号帳励振源用利得部をそれぞれ分割し、ピッチ周期成分と符号帳の励振ベクトル成分それを半分ずつピッチ励振源用利得部および符号帳励振源用利得部に与えることを特徴とする音声の予測符号化装置。

【請求項4】 請求項2に記載される音声の予測符号化装置において、利得部をピッチ励振源用利得部と符号帳励振源用利得部とを縦続接続して構造化したことを特徴とする音声の予測符号化装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】この発明は、音声の予測符号化方法および装置に関し、特に、2個の励振源からそれぞれ選択した励振候補により合成フィルタを駆動してより少ない情報量の音声信号系列ディジタル伝送を実現する音声の予測符号化方法および装置に関する。

【0002】

【従来の技術】ディジタル移動通信その他の音声の情報

量、伝送周波数帯域を圧縮して伝送する技術を採用する技術分野においては、種々の高能率符号化法が採用されている。8 kbit/s程度の情報量により符号化する方法としては、CELP(符号駆動型線形予測)、VSELP(ベクトル加算駆動型線形予測)、マルチパルス符号化その他の方法が知られている。これらの技術は、M.R.Schroeder and B.S.Atal : "Code-Excited Linear Prediction (CELP): High-quality Speech at Very Low Rates", Proc. ICASSP'85, 25.1.1, pp. 937-940, 1985、或はI.

- 10 A.Gerson and M.A.Jasiuk : "VectorSum Excited Linear Prediction (VSELP) Speech Coding at 8 kps", proc. ICASSP'90, S9.3, pp. 461-464, 1990 或は、小澤一範、荒関 隼 ; "ピッチ情報を用いた9.6-4.8kbit/s マルチパルス音声符号化方式", 信学誌(D-II), J72-D-II, 8, pp. 1125-1132, 1989 その他の文献を開示している。

【0003】これらの方程式は、図1に示される通り、入力音声を複数サンプリングしたものをフィルタ係数決定部10に供給して、ここにおいてフィルタ係数を計算する。計算されたフィルタ係数はフィルタ係数量子化部11に供給されて、ここにおいて量子化される。量子化されたフィルタ係数を合成フィルタ12に設定する。A(z)は合成フィルタ12の伝達関数である。ピッチ励振源13の複数のピッチ周期成分(励振候補)から選択されたピッチ周期成分と、符号帳励振源14の複数の雑音波形ベクトル成分(例えば乱数ベクトル成分、励振候補)から選択された候補とに対してそれぞれ利得部15および利得部16において適当な利得を加えた後、これらを加算器において加算したものを駆動信号として合成フィルタ12に供給し、音声を合成する。入力音声パワー量子化部19は入力音声のパワーを計算し、これを量子化して入力音声とピッチ周期成分を正規化する。合成音声と正規化された入力音声との間の差である歪が最も小さくなるように歪パワー計算部17において両励振源中の各励振候補を選択し、かつ利得部15および利得部16の各利得を設定する。符号出力部18は予測係数、入力音声パワー、ピッチ周期成分候補と符号帳の候補それぞれに対して選ばれたコード番号および利得を符号として出力する。これらの方程式は、上述の通り、入力音声のパワーを量子化して伝送するものである。この量子化されたパワー情報が伝送路で誤った場合、復号音声の品質は著しく劣化する。

- 【0004】一方、パワー情報を使用しない方式として、16 kbit/s程度の情報量を使用したLD-CELP(低遅延符号駆動型線形予測)符号化方式がある(J.Chen ; "High Quality 16 kb/s Speech Coding with a One-Way Delay Less Than 2 ms", Proc. ICASSP'90, 33.s 9.1, 1990. 参照)。この方法はパワーの後方予測を採用している。図2に示される如く、現在量子化しようとするフレーム内のパワーを量子化せず、利得予測部23に

において過去の残差波形のパワーから線形予測により次のフレームのパワー（即ち、利得）を予測している。また、復号波形をフィルタ係数決定部25において50次の高次の予測を行い、ピッチの周期性も含めた線形予測を行う。合成フィルタ24にフィルタ係数をセットし、その合成フィルタ $1/A(z)$ を使用して符号帳励振源中の形状ベクトル成分（雑音成分）候補を求め、その符号を伝送する。この方法は符号器と復号器の双方において過去に復号化された音声を共通に利用することができる、音声パワーの情報を伝送する必要がない。

【0005】しかし、LD-CELPは現在のフレームの予測を過去の復号化された系列のみから行うものであるので、予測誤差が従来の前方予測型と比較して大きい。そのため、8 kbit/s程度の情報量による符号化では急激に波形歪が増大し、品質が劣化する。従って、8 kbit/s程度の情報量により高品質の音声符号化を実現するには、LD-CELPの如くピッチの周期性を線形予測に含めるのではなく、ピッチ周期成分も復号化された音声から抽出する必要がある。

【0006】

【発明が解決しようとする課題】この発明は、雑音符号帳を使用する駆動型線形予測符号化方法および装置において、伝送路誤りが発生したときに復号音声に聽感上の大変な劣化を与える音声利得に関する情報を伝送することなく、良い品質の音声を復号することができる予測符号化方法および装置を提供するものである。

【0007】

【課題を解決するための手段】入力音声を複数サンプリングしたものに基づいて計算された量子化フィルタ係数の設定された合成フィルタを2個の励振源から時系列ベクトル成分によりフレーム単位毎に駆動して音声を再生することを利用して音声を符号化する音声の予測符号化方法において、2個の励振源のベクトル成分に対してそれぞれの利得を同時に求め、2個の励振源の内的一方である複数の雑音波形ベクトル成分を有する符号帳励振源31の選択された雑音波形ベクトル成分に選択された利得を乗じた残差波形とそれ以前に蓄積されている残差波形とに基づいて次のフレームの符号帳励振源31の雑音波形ベクトル成分に与える利得を予測することにより、パワーの情報を伝送することなく音声を符号化する音声の予測符号化方法を構成し、そして、入力音声を複数サンプリングしたものに基づいて計算された量子化フィルタ係数の設定された合成フィルタ37を具備し、複数のピッチ周期成分を有するピッチ励振源30と複数の雑音波形ベクトル成分を有する符号帳励振源31とよりなる2個の励振源からフレーム単位毎に時系列ベクトル成分により合成フィルタ37を駆動して音声を再生することを利用して音声を符号化する音声の予測符号化装置において、2個の励振源30および31のベクトル成分それに対して利得を同時に乗算する利得部33を具備

し、合成音声と入力音声との間の差である歪を最小にする様に両励振源中の各励振ベクトル成分を選択すると共に利得部33の利得を設定する歪パワー計算部38を具備し、符号帳励振源31の選択された雑音波形ベクトル成分に選択された利得を乗じた残差波形とそれ以前に蓄積されている残差波形とに基づいて次のフレームの符号帳の雑音波形ベクトル成分に与える利得を予測する利得予測部34を具備する音声の予測符号化装置を構成し、また、利得部42はピッチ励振源用利得部421および符号帳励振源用利得部422をそれぞれ分割し、ピッチ周期成分と符号帳の励振ベクトル成分それを半分ずつピッチ励振源用利得部421および符号帳励振源用利得部422に与える音声の予測符号化装置を構成し、更に、利得部33をピッチ励振源用利得部50と符号帳励振源利得部51とを継続接続して構造化した音声の予測符号化装置をも構成した。

【0008】

【実施例】この発明の実施例を図3を参照して説明する。入力音声を複数サンプリングしたものをフィルタ係

20 数決定部35に供給し、ここにおいてフィルタ係数を計算する。計算されたフィルタ係数は次いでフィルタ係数量子化部36に供給され、フィルタ係数はここにおいて量子化されて、その量子化されたフィルタ係数は合成フィルタ37に設定される。A(z)は合成フィルタ37の伝達関数である。合成フィルタ37の励振源は2個よりなる。その一方はピッチ励振源30の複数のピッチ周期成分（励振候補）から選択されたピッチ周期成分に利得部33において適当な利得を乗じた成分であり、他方は符号帳励振源31の複数の雑音波形ベクトル成分（例えれば乱数ベクトル成分、励振候補）から選択された候補であって予測利得部32において予測利得倍した候補に利得部33において適当な利得を乗じた成分である。ここで、利得部33における選択されたピッチ周期成分および雑音波形ベクトル成分に対する利得の乗算は符号化効率を上げるために同時に実施される。歪パワー計算部38は合成音声と入力音声との間の差である歪が最も小さくなる様に両励振源中の各励振候補を選択し、かつ利得部33の利得を設定する。符号出力部39はフィルタ係数、ピッチ周期成分候補および符号帳の候補のそれぞれについて、選択されたコード番号および利得を符号として出力する。利得予測部34は、符号帳励振源31の選択された候補を予測利得部32において予測利得倍したものに利得部33において選択された利得を乗じた残差波形とそれ以前に蓄積されている残差波形とに基づいて、次のフレームの符号帳の候補に与える利得（パワー）を予測する。

【0009】ところで、利得の情報が伝送路において誤った場合も復号音声はその影響を蒙る。この影響を少なくするために図3における利得部33を分割する。これを図4を参照して説明するに、利得部42はピッチ励振

源用利得部421および符号帳励振源用利得部422をそれぞれ分割し、ピッチ周期成分と符号帳の励振ベクトル成分それを半分ずつピッチ励振源用利得部421および符号帳励振源用利得部422に与える。全体として合成音声と入力音声との間の差である歪が小さくなる様に個々の利得はピッチ周期成分と励振ベクトル成分成分を一括して決定する。この様に利得部を分割することにより、何れか一方の利得情報が伝送路において誤っても、他方の利得情報によってこれを補うことができ、伝送路誤りによる影響を少なくすることができる。

【0010】また、利得の情報が伝送路において誤った場合の対策として、図3における利得部33を図5に示される如くに、ピッチ励振源用利得部50と符号帳励振源用利得部51とを継続接続して構造化する。ピッチ周期成分は $\{g_{p0} \sim g_{pM}\}$ の中から選択される。符号帳の励振ベクトル成分は、例えばピッチ周期成分 g_{p0} が選択された場合は、 $\{g_{c0} \sim g_{cN}\}$ の中から選択される。このように構造化しておけば、符号帳の励振ベクトル成分の利得情報が誤ってもピッチ周期成分の情報はまったく影響を受けない。逆に、ピッチ周期成分の利得情報が誤った場合は、符号帳の利得情報も誤るが、符号帳の利得のコードを適切に並べておくことによりその影響を少なくすることができる。

【0011】

【発明の効果】従来の音声予測符号化方法および装置法は、先ず入力音声のパワーを計算し、そのパワーを量子化する。この量子化されたパワーを用いて入力音声とピッチ成分の正規化を行い、利得を決定する。そのためには量子化されたパワーに関する情報を伝送する必要があ *

*る。この発明の音声予測符号化方法および装置は、音声パワーの情報を伝送することなく、過去の符号帳の励振ベクトル成分に基づいて次のフレームの利得（音声パワー）を予測して符号化することを特徴とするものであるので、伝送路においてビット誤りが発生しても音声のパワー情報を伝送していないために、復号音声のパワーはこのビット誤りによる大きな影響を蒙ることなく、高品質の音声の符号化および復号化を実現することができる。

10 【図面の簡単な説明】

【図1】従来の音声予測符号化装置のブロック図。

【図2】従来の音声予測符号化装置であるLD-CEL Pのブロック図。

【図3】この発明の音声予測符号化装置のブロック図。

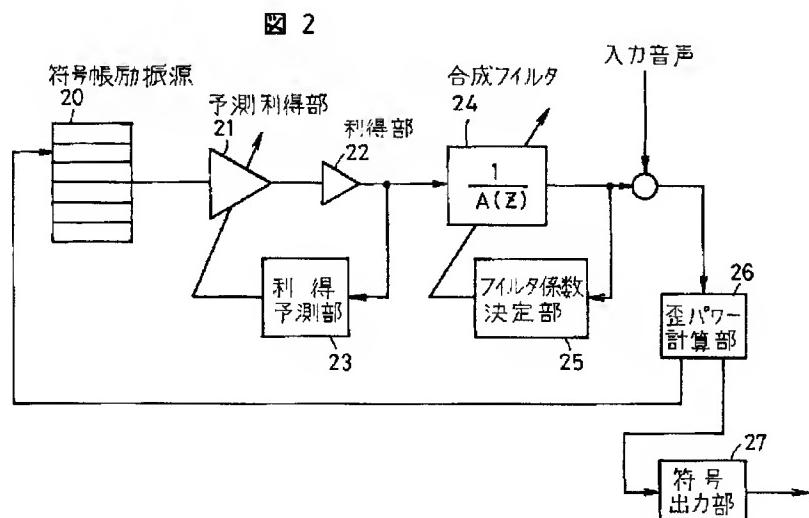
【図4】分割した利得部のブロック図。

【図5】構造化した利得部のブロック図。

【符号の説明】

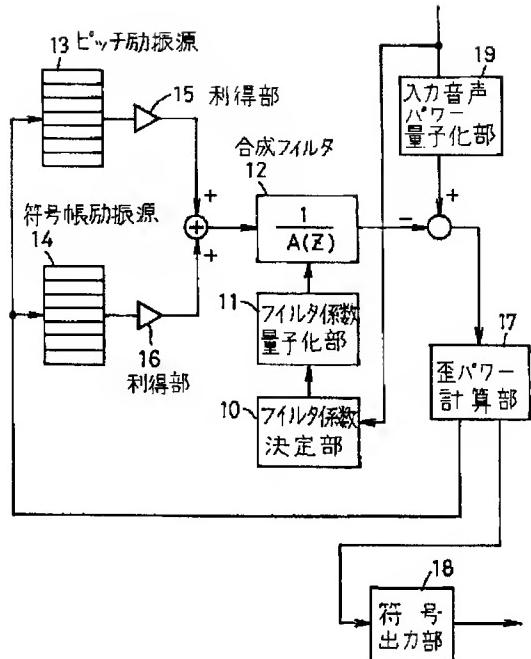
30	ピッチ励振源
31	符号帳励振源
20	利得部
33	利得予測部
34	利得部
37	合成フィルタ
38	歪パワー計算部
42	利得部
421	ピッチ励振源用利得部
422	符号帳励振源用利得部
50	ピッチ励振源用利得部
51	符号帳励振源用利得部

【図2】



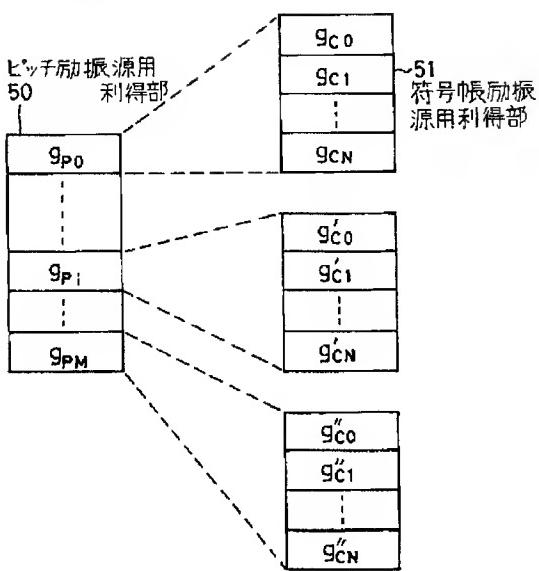
【図1】

図1



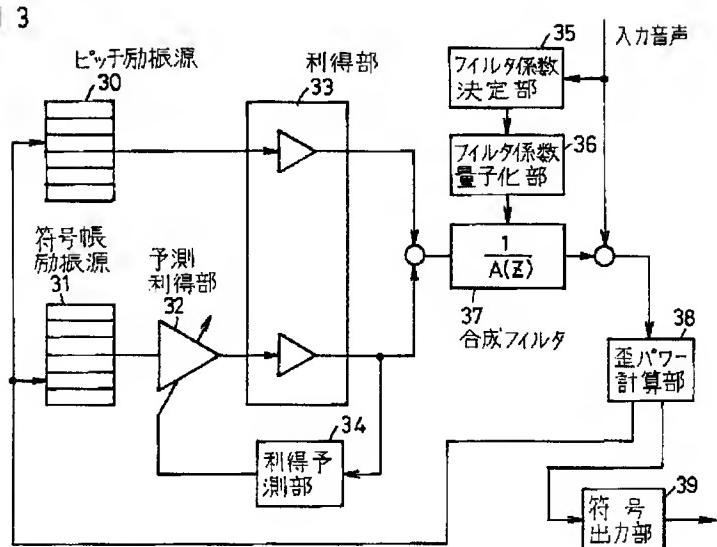
【図5】

図5



【図3】

図3



【図4】

